



OICE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/885,188

DATE: 01/25/2002

TIME: 17:01:04

Input Set : N:\Crf3\RULE60\09885188.raw

Output Set: N:\CRF3\01252002\I885188.raw

## SEQUENCE LISTING

## 1 (1) GENERAL INFORMATION:

2 (i) APPLICANT: Chris Somerville

3 Pierre Broun

4 Frank van de Loo

5 (ii) TITLE OF INVENTION: Production of Hydroxylated Fatty Acids in

6 Genetically Modified Plants

7 (iii) NUMBER OF SEQUENCES: 15

8 (iv) CORRESPONDENCE ADDRESS:

9 (A) ADDRESSEE: Pillsbury Winthrop, L.L.P.

10 (B) STREET: 1600 Tysons Boulevard

11 (C) CITY: McLean

12 (D) STATE: VA

13 (E) COUNTRY: USA

14 (F) ZIP: 22102

15 (v) COMPUTER READABLE FORM:

16 (A) MEDIUM TYPE: Diskette, 3.50 inch

17 (B) COMPUTER: IBM PC-compatible

18 (C) OPERATING SYSTEM: MS-DOS

19 (D) SOFTWARE: MS Word

20 (vi) CURRENT APPLICATION DATA:

C--&gt; 21 (A) APPLICATION NUMBER: US/09/885,188

C--&gt; 22 (B) FILING DATE: 21-Jun-2001

23 (C) CLASSIFICATION:

24 (vii) PRIOR APPLICATION DATA:

25 (A) APPLICATION NUMBER: US/08/530,862B

26 (B) FILING DATE: 06-Feb-1996

27 (A) APPLICATION NUMBER: PCT/US95/11855

28 (B) FILING DATE: September 25, 1995

29 (A) APPLICATION NUMBER: US 08/530,862

30 (B) FILING DATE: September 20, 1995

31 (A) APPLICATION NUMBER: US 08/320,982

32 (B) FILING DATE: October 11, 1994

33 (A) APPLICATION NUMBER: US 08/314,596

34 (B) FILING DATE: September 26, 1994

35 (2) INFORMATION FOR SEQ ID NO: 1

36 (i) SEQUENCE CHARACTERISTICS:

37 (A) LENGTH: 543 nucleotides

38 (B) TYPE: nucleotide

39 (C) STRANDEDNESS: single

40 (D) TOPOLOGY: linear

41 (ii) MOLECULE TYPE: cDNA

42 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/885,188

DATE: 01/25/2002

TIME: 17:01:04

Input Set : N:\Crif3\RULE60\09885188.raw

Output Set: N:\CRF3\01252002\I885188.raw

|    |   |     |
|----|---|-----|
| 43 | TATTGGCACC GGCGGCACCA TTCCAACAAT GGATCCCTAG AAAAAGATGA AGTCTTTGTC | 60  |
| 44 | CCACCTAAGA AAGCTGCAGT CANATGGTAT GTCAAATACC TCAACAACCC TCTTGGACGC | 120 |
| 45 | ATTCTGGTGT TAACAGTTCA GTTTATCCTC GGGTGGCCTT TGTATCTAGC CTTTAATGTA | 180 |
| 46 | TCAGGTAGAC CTTATGATGG TTTCGCTTCA CATTTCTTCC CTCATGCACC TATCTTTAAG | 240 |
| 47 | GACCGTGAAC GTCTCCAGAT ATACATCTCA GATGCTGGTA TTCTAGCTGT CTGTTATGGT | 300 |
| 48 | CTTTACCGTT ACGCTGCTTC ACAAGGATTG ACTGCTATGA TCTGCGTCTA CGGAGTACCG | 360 |
| 49 | CTTTTGATAG TGAACCTTTT CCTTGCTTGT GTCACCTTCT TGCAGCACAC TCATCCTTCA | 420 |
| 50 | TTACCTCACT ATGATTCAAC CGAGTGGGAA TGGATTAGAG GAGCTTTGGT TACGGTAGAC | 480 |
| 51 | AGAGACTATG GAATCTTGAA CAAGGTGTTT CACAACATAA CAGACACCCA CGTAGCACAC | 540 |
| 52 | CAC   | 543 |

## 54 (2) INFORMATION FOR SEQ ID NO: 2

## 55 (i) SEQUENCE CHARACTERISTICS:

56 (A) LENGTH: 544 nucleotides

57 (B) TYPE: nucleotide

58 (C) STRANDEDNESS: single

59 (D) TOPOLOGY: linear

## 60 (ii) MOLECULE TYPE: cDNA

## 61 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

|    |   |     |
|----|---|-----|
| 62 | TATAGGCACC GGAGGCACCA TTCCAACACA GGATCCCTCG AAAGAGATGA AGTATTTGTC | 60  |
| 63 | CCAAAGCAGA AATCCGCAAT CAAGTGGTAC GGCGAATACC TCAACAACCC TCCTGGTCGC | 120 |
| 64 | ATCATGATGT TAACTGTCCA GTTCGTCCTC GGATGGCCCT TGTACTTAGC CTTCAACGTT | 180 |
| 65 | TCTGGCAGAC CCTACAATGG TTTCGCTTCC CATTTCTTCC CCAATGCTCC TATCTACAAC | 240 |
| 66 | GACCGTGAAC GCCTCCAGAT TTACATCTCT GATGCTGGTA TTCTAGCCGT CTGTTATGGT | 300 |
| 67 | CTTTACCGTT ACGCTGTTGC ACAAGGACTA GCCTCAATGA TCTGTCTAAA CGGAGTTCCG | 360 |
| 68 | CTTCTGATAG TTAACCTTTT CCTCGTCTTG ATCACTTACT TACAACACAC TCACCCTGCG | 420 |
| 69 | TTGCCTCACT ATGATTCATC AGAGTGGGAT TGGCTTAGAG GAGCTTTAGC TACTGTAGAC | 480 |
| 70 | AGAGACTATG GAATCTTGAA CAAGGTGTTC CATAACATCA CAGACACCCA CGTCGCACAC | 540 |
| 71 | CACT  | 544 |

## 73 (2) INFORMATION FOR SEQ ID NO: 3

## 74 (i) SEQUENCE CHARACTERISTICS:

75 (A) LENGTH: 1855 nucleotides

76 (B) TYPE: nucleotide

77 (C) STRANDEDNESS: single

78 (D) TOPOLOGY: linear

## 79 (ii) MOLECULE TYPE: genomic

## 80 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

|    |  |     |
|----|--|-----|
| 81 | ATGAAGCTTT ATAAGAAGTT AGTTTTCTCT GGTGACAGAG AAATTNTGTC AATTGGTAGT  | 60  |
| 82 | GACAGTTGAA GCAACAGGAA CAACAAGGAT GGTGAGTGNT GATGCTGATG TGGTGATGTG  | 120 |
| 83 | TTATTCATCA AATACTAAAT ACTACATTAC TTGTTGCTGC CTACTTCTCC TATTTCTCTC  | 180 |
| 84 | GCCACCCATT TTGGACCCAC GANCTTCCA TTTAAACCCT CTCTCGTGCT ATTCACCAGA   | 240 |
| 85 | AGAGAAGCCA AGAGAGAGAG AGAGAGAATG TTCTGAGGAT CATTGTCTTC TTCATCGTTA  | 300 |
| 86 | TTAACGTAAG TTTTTTTTGA CCACTCATAT CTAAAATCTA GTACATGCAA TAGATTAATG  | 360 |
| 87 | ACTGTTCCCTT CTTTTGATAT TTTCAGCTTC TTGAATTCAA GATGGGTGCT GGTGGAAGAA | 420 |
| 88 | TAATGGTTAC CCCCTCTTCC AAGAAATCAG AAAGTGAAGC CCTAAAACGT GGACCATGTG  | 480 |
| 89 | AGAAACCACC ATTCAGTGT AAAGATCTGA AGAAAGCAAT CCCACAGCAT TGTTTCAAGC   | 540 |
| 90 | GCTCTATCCC TCGTTCTTTC TCCTACCTTC TCACAGATAT CACTTTAGTT TCTTGCTTCT  | 600 |
| 91 | ACTACGTTGC CACAAATTAC TTCTCTCTTC TTCCTCAGCC TCTCTCTACT TACCTAGCTT  | 660 |
| 92 | GGCCTCTCTA TTGGGTATGT CAAGGCTGTG TCTTAACCGG TATCTGGGTC ATTGGCCATG  | 720 |
| 93 | AATGTGGTCA CCATGCATTC AGTGACTATC AATGGGTAGA TGACACTGTT GGTTTTATCT  | 780 |

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/885,188

DATE: 01/25/2002

TIME: 17:01:04

Input Set : N:\Crif3\RULE60\09885188.raw

Output Set: N:\CRF3\01252002\I885188.raw

```

94   TCCATTCCTT CCTTCTCGTC CCTTACTTCT CCTGGAAATA CAGTCATCGT CGTCACCATT      840
95   CCAACAATGG ATCTCTCGAG AAAGATGAAG TCTTTGTCCC ACCGAAGAAA GCTGCAGTCA      900
96   AATGGTATGT TAAATACCTC AACAACCCTC TTGGACGCAT TCTGGTGTTA ACAGTTCAGT      960
97   TTATCCTCGG GTGGCCTTTG TATCTAGCCT TTAATGTATC AGGTAGACCT TATGATGGTT     1020
98   TCGCTTCACA TTTCTTCCCT CATGCACCTA TCTTTAAAGA CCGAGAACGC CTCCAGATAT     1080
99   ACATCTCAGA TGCTGGTATT CTAGCTGTCT GTTATGGTCT TTACCGTTAC GCTGCTTCAC     1140
100  AAGGATTGAC TGCTATGATC TGCGTCTATG GAGTACCGCT TTTGATAGTG AACTTTTTC      1200
101  TTGTCTTGGT AACTTTCTTG CAGCACACTC ATCCTTCGTT ACCTCATTAT GATTCAACCG      1260
102  AGTGGGAATG GATTAGAGGA GCTTTGGTTA CGGTAGACAG AGACTATGGA ATATTGAACA      1320
103  AGGTGTTCCA TAACATAACA GACACACATG TGGCTCATCA TCTCTTTGCA ACTATACCGC      1380
104  ATTATAACGC AATGGAAGCT ACAGAGGCCA TAAAGCCAAT ACTTGGTGAT TACTACCACT      1440
105  TCGATGGAAC ACCGTGGTAT GTGGCCATGT ATAGGGAAGC AAAGGAGTGT CTCTATGTAG      1500
106  AACCGGATAC GGAACGTGGG AAGAAAGGTG TCTACTATTA CAACAATAAG TTATGAGGCT      1560
107  GATAGGGCGA GAGAAGTGCA ATTATCAATC TTCATTTCCA TGTTTTAGGT GTCTTGTTTA      1620
108  AGAAGCTATG CTTTGTTTCA ATAATCTCAG AGTCCATNTA GTTGTGTTCT GGTGCATTTT      1680
109  GCCTAGTTAT GTGGTGTCTG AAGTTAGTGT TCAAAC TGCT TCCTGCTGTG CTGCCCAGTG      1740
110  AAGAACAAGT TTACGTGTTT AAAATACTCG GAACGAATTG ACCACAANAT ATCCAAAACC      1800
111  GGCTATCCGA ATTCCATATC CGAAAACCGG ATATCCAAAT TTCCAGAGTA CTTAG          1855

```

113 (2) INFORMATION FOR SEQ ID NO: 4

114 (i) SEQUENCE CHARACTERISTICS:

115 (A) LENGTH: 384 amino acids

116 (B) TYPE: amino acid

117 (C) STRANDEDNESS:

118 (D) TOPOLOGY: linear

119 (ii) MOLECULE TYPE: protein

120 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

121  Met Gly Ala Gly Gly Arg Ile Met Val Thr Pro Ser Ser Lys Lys Ser
122      1              5              10              15
123  Glu Thr Glu Ala Leu Lys Arg Gly Pro Cys Glu Lys Pro Pro Phe Thr
124      20              25              30
125  Val Lys Asp Leu Lys Lys Ala Ile Pro Gln His Cys Phe Lys Arg Ser
126      35              40              45
127  Ile Pro Arg Ser Phe Ser Tyr Leu Leu Thr Asp Ile Thr Leu Val Ser
128      50              55              60
129  Cys Phe Tyr Tyr Val Ala Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro
130      65              70              75              80
131  Leu Ser Thr Tyr Leu Ala Trp Pro Leu Tyr Trp Val Cys Gln Gly Cys
132      85              90              95
133  Val Leu Thr Gly Ile Trp Val Ile Gly His Glu Cys Gly His His Ala
134      100             105             110
135  Phe Ser Asp Tyr Gln Trp Val Asp Asp Thr Val Gly Phe Ile Phe His
136      115             120             125
137  Ser Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg
138      130             135             140
139  His His Ser Asn Asn Gly Ser Leu Glu Lys Asp Glu Val Phe Val Pro
140      145             150             155             160
141  Pro Lys Lys Ala Ala Val Lys Trp Tyr Val Lys Tyr Leu Asn Asn Pro
142      165             170             175
143  Leu Gly Arg Ile Leu Val Leu Thr Val Gln Phe Ile Leu Gly Trp Pro

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/885,188

DATE: 01/25/2002

TIME: 17:01:05

Input Set : N:\Crif3\RULE60\09885188.raw

Output Set: N:\CRF3\01252002\I885188.raw

|     |   |   |     |     |     |     |
|-----|---|---|-----|-----|-----|-----|
| 144 |   | 180   |     | 185 |     | 190 |
| 145 | Leu Tyr   | Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Phe Ala |     |     |     |     |
| 146 |   | 195   |     | 200 |     | 205 |
| 147 | Ser His Phe Phe Pro His Ala Pro Ile Phe Lys Asp Arg Glu Arg Leu |   |     |     |     |     |
| 148 |   | 210   |     | 215 |     | 220 |
| 149 | Gln Ile Tyr Ile Ser Asp Ala Gly Ile Leu Ala Val Cys Tyr Gly Leu |   |     |     |     |     |
| 150 |   | 225   |     | 230 |     | 235 |
| 151 | Tyr Arg Tyr Ala Ala Ser Gln Gly Leu Thr Ala Met Ile Cys Val Tyr |   |     |     |     |     |
| 152 |   |   | 245 |     | 250 | 255 |
| 153 | Gly Val Pro Leu Leu Ile Val Asn Phe Phe Leu Val Leu Val Thr Phe |   |     |     |     |     |
| 154 |   | 260   |     | 265 |     | 270 |
| 155 | Leu Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Thr Glu Trp |   |     |     |     |     |
| 156 |   | 275   |     | 280 |     | 285 |
| 157 | Glu Trp Ile Arg Gly Ala Leu Val Thr Val Asp Arg Asp Tyr Gly Ile |   |     |     |     |     |
| 158 |   | 290   |     | 295 |     | 300 |
| 159 | Leu Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His |   |     |     |     |     |
| 160 |   | 305   |     | 310 |     | 315 |
| 161 | Leu Phe Ala Thr Ile Pro His Tyr Asn Ala Met Glu Ala Thr Glu Ala |   |     |     |     |     |
| 162 |   |   | 325 |     | 330 | 335 |
| 163 | Ile Lys Pro Ile Leu Gly Asp Tyr Tyr His Phe Asp Gly Thr Pro Trp |   |     |     |     |     |
| 164 |   | 340   |     | 345 |     | 350 |
| 165 | Tyr Val Ala Met Tyr Arg Glu Ala Lys Glu Cys Leu Tyr Val Glu Pro |   |     |     |     |     |
| 166 |   | 355   |     | 360 |     | 365 |
| 167 | Asp Thr Glu Arg Gly Lys Lys Gly Val Tyr Tyr Tyr Asn Asn Lys Leu |   |     |     |     |     |
| 168 |   | 370   |     | 375 |     | 380 |

170 (2) INFORMATION FOR SEQ ID NO: 5

171 (i) SEQUENCE CHARACTERISTICS:

172 (A) LENGTH: 387 amino acids

173 (B) TYPE: amino acid

174 (C) STRANDEDNESS:

175 (D) TOPOLOGY: linear

176 (ii) MOLECULE TYPE: protein

177 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

|     |   |
|-----|---|
| 178 | Met Gly Gly Gly Gly Arg Met Ser Thr Val Ile Thr Ser Asn Asn Ser |
| 179 | 1 5 10 15   |
| 180 | Glu Lys Lys Gly Gly Ser Ser His Leu Lys Arg Ala Pro His Thr Lys |
| 181 | 20 25 30  |
| 182 | Pro Pro Phe Thr Leu Gly Asp Leu Lys Arg Ala Ile Pro Pro His Cys |
| 183 | 35 40 45  |
| 184 | Phe Glu Arg Ser Phe Val Arg Ser Phe Ser Tyr Val Ala Tyr Asp Val |
| 185 | 50 55 60  |
| 186 | Cys Leu Ser Phe Leu Phe Tyr Ser Ile Ala Thr Asn Phe Phe Pro Tyr |
| 187 | 65 70 75 80   |
| 188 | Ile Ser Ser Pro Leu Ser Tyr Val Ala Trp Leu Val Tyr Trp Leu Phe |
| 189 | 85 90 95  |
| 190 | Gln Gly Cys Ile Leu Thr Gly Leu Trp Val Ile Gly His Glu Cys Gly |
| 191 | 100 105 110   |
| 192 | His His Ala Phe Ser Glu Tyr Gln Leu Ala Asp Asp Ile Val Gly Leu |
| 193 | 115 120 125   |

## RAW SEQUENCE LISTING

DATE: 01/25/2002

PATENT APPLICATION: US/09/885,188

TIME: 17:01:05

Input Set : N:\Crif3\RULE60\09885188.raw

Output Set: N:\CRF3\01252002\I885188.raw

```

194   Ile Val His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser
195       130                               135               140
196   His Arg Arg His His Ser Asn Ile Gly Ser Leu Glu Arg Asp Glu Val
197   145                               150               155               160
198   Phe Val Pro Lys Ser Lys Ser Lys Ile Ser Trp Tyr Ser Lys Tyr Ser
199       165                               170               175
200   Asn Asn Pro Pro Gly Arg Val Leu Thr Leu Ala Ala Thr Leu Leu Leu
201       180                               185               190
202   Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp
203       195                               200               205
204   Arg Phe Ala Cys His Tyr Asp Pro Tyr Gly Pro Ile Phe Ser Glu Arg
205       210                               215               220
206   Glu Arg Leu Gln Ile Tyr Ile Ala Asp Leu Gly Ile Phe Ala Thr Thr
207   225                               230               235               240
208   Phe Val Leu Tyr Gln Ala Thr Met Ala Lys Gly Leu Ala Trp Val Met
209       245                               250               255
210   Arg Ile Tyr Gly Val Pro Leu Leu Ile Val Asn Cys Phe Leu Val Met
211       260                               265               270
212   Ile Thr Tyr Leu Gln His Thr His Pro Ala Ile Pro Arg Tyr Gly Ser
213       275                               280               285
214   Ser Glu Trp Asp Trp Leu Arg Gly Ala Met Val Thr Val Asp Arg Asp
215       290                               295               300
216   Tyr Gly Val Leu Asn Lys Val Phe His Asn Ile Ala Asp Thr His Val
217   305                               310               315               320
218   Ala His His Leu Phe Ala Thr Val Pro His Tyr His Ala Met Glu Ala
219       325                               330               335
220   Thr Lys Ala Ile Lys Pro Ile Met Gly Glu Tyr Tyr Arg Tyr Asp Gly
221       340                               345               350
222   Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Lys Glu Cys Leu Phe
223       355                               360               365
224   Val Glu Pro Asp Glu Gly Ala Pro Thr Gln Gly Val Phe Trp Tyr Arg
225       370                               375               380
226   Asn Lys Tyr
227       385

```

229 (2) INFORMATION FOR SEQ ID NO: 6

230 (i) SEQUENCE CHARACTERISTICS:

231 (A) LENGTH: 383 amino acids

232 (B) TYPE: amino acid

233 (C) STRANDEDNESS:

234 (D) TOPOLOGY: linear

235 (ii) MOLECULE TYPE: protein

236 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```

237   Met Gly Ala Gly Gly Arg Met Pro Val Pro Thr Ser Ser Lys Lys Ser
238       1                               5               10               15
239   Glu Thr Asp Thr Thr Lys Arg Val Pro Cys Glu Lys Pro Pro Phe Ser
240       20                               25               30
241   Val Gly Asp Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser
242       35                               40               45
243   Ile Pro Arg Ser Phe Ser Tyr Leu Ile Ser Asp Ile Ile Ile Ala Ser

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/885,188

DATE: 01/25/2002

TIME: 17:01:06

Input Set : N:\Crif3\RULE60\09885188.raw

Output Set: N:\CRF3\01252002\I885188.raw

L:21 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:22 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]